A Mixture Model for NLOS mmWave Interference Distribution

Hussain Elkotby and Mai Vu
Department of Electrical and Computer Engineering, Tufts University, Medford, MA, USA
Emails: Hussain.Elkotby@Tufts.edu, Mai.Vu@Tufts.edu

***Abstract*—We propose a novel model for the NLOS interference power distribution in a cellular system employing MIMO beamforming transmission in mmWave spectrum. The proposed model is a mixture of the Inverse Gaussian and the Inverse Weibull distributions, both are two-parameter medium-to-heavy tail distributions. To estimate the parameters of the mixture model, we design an expectation maximization algorithm using both analytical moment matching and maximum likelihood estimation. Further, the information-theoretic metric relative entropy is used to measure the goodness of the model by capturing the distance from the modeled distribution to a reference one. The model is tested against simulation data obtained from a stochastic geometry based cellular network and demonstrates very good fit for a wide range of practical mmWave path loss and shadowing values. The mixture model is then used to analyze the capacity of a cellular network employing joint dominant mode beamforming and again confirms a good fit with simulation.**

Keywords: mmWave cellular; interference model; moment matching; maximum likelihood estimation; mixture distribution.

I. Introduction

The current scarcity of wireless spectrum coupled with the predicted exponential increase in capacity demand has focused attention on the underutilized mmWave bands (3-300 GHz) [1]–[3]. mmWave makes a promising candidate for 5G systems through the huge amount of available spectrum [2], [4]. Traditionally, mmWave bands are considered for backhaul in cellular systems, but not for cellular access [2], [4]. Recent trials, however, show that nodes equipped with large antenna array can successfully compensate for the high path loss and establish high-rate communication over short distances (100-200m), even in NLOS environments [1], [2]. As a result, mmWave is being considered for cellular access networks with short cell radius.

Smaller cell radii lead to dense base station (BS) deployments. Even under beamforming, these high BS and user densities can drive the cellular networks to be more interference rather than noise limited. While many-element adaptive arrays can boost the received signal power and hence reduce the impact of out-of-cell interference [3], [5], interference characterization still plays an important role in evaluating and predicting the dense mmWave networks performance.

For cellular network analysis, stochastic geometry has been shown to be analytically tractable and capture the main performance trends. In a recent work, the Gaussian, Gamma, Inverse Gamma, and Inverse Gaussian distributions using moment matching are examined for modeling the interference distribution [6]. Both the Gamma and Inverse Gaussian are shown to be suitable for the Poisson point process network. This prior work, however, considered only Rayleigh fading with no shadowing.

mmWave cellular systems differ considerably from conventional systems due to the particular channel characteristics. Not only the propagation path loss is much higher with distance, but measurement results show that mmWave sensitivity to blockages also leads to severe shadowing effects. Shadowing is of particular concerns in urban environments where mmWave is predicted to have wide usage [7].

In our work, we find that shadowing introduces large variation in the interference power and causes a medium to heavy tail, therefore we introduce a new interference power model as a mixture of the Inverse Gaussian (IG) and Inverse Weibull (IW), which are heavy tail distributions that can capture shadowing with large variance. We consider a combination of moment matching (MM) and maximum likelihood estimation (MLE) to estimate the mixture model parameters. In our model development, we consider the interference at a BS in a single user MIMO beamforming in an uplink scenario, but the result is also applicable to the downlink. The model applies to all transmit beamforming techniques, as long as the beamforming vector is independent of the interference channels.

II. Models for network and channel propagation

A. Network model based on stochastic geometry

We consider a cellular system consisting of multiple cells with an average cell radius R_0. Each cell has a single BS that is equipped with N_{BS} antennas and serves multiple user equipments (UEs). Each UE is equipped with N_{UE} antennas and uses a distinct resource block in each cell, hence there is no intra-cell interference. However, each UE suffers from out-of-cell interference due to frequency reuse in all other cells. In this paper, we denote all transmissions on the same resource block from all cells other than the cell under study as the out-of-cell interference. Further, we assume that each UE is served by the single BS closest to that user. Extensions to other cell association schemes such as those based on the strongest received power can be considered in future work.

Due to the irregular structure of current and future cellular networks, we employ stochastic geometry to describe the network. We consider uplink transmissions in this paper. We model the active UEs in different cells contending for the same resource block and causing interference to each other as being distributed on a two-dimensional plane according to a homogeneous and stationary Poisson point process (p.p.p.) Φ_1 with intensity λ_1.

In order to develop a model for the out-of-cell interference power in a stochastic geometry based network, we fix the radius of the cell under study to be R_c, which is typically inversely proportional to the active UEs density as $R_c = 1/(2\sqrt{\lambda_1/\eta})$, where η represents the user density factor and can be varied. Further, we consider single user MIMO beamforming, but our results can be directly applied to multi-user MIMO (MU-MIMO) by simply increasing the interferers density. As such, our network model consists of a cell under study centered at its BS and surrounded by Voronoi cells of other active uplink UEs who interfere with the considered BS at the origin. Our goal is to model this out-of-cell interference to the considered BS receiver, taking into account mmWave channel propagation as discussed next.

B. mmWave channel propagation model

In our analysis, we consider a complete channel model with shadowing, path loss and small scale fading. We also employ the simplified step function approximation of LOS and NLOS regions introduced in [8] and assume that the interferers are located outside a cell of radius R_c that is larger than the radius of the LOS ball region, such that interference only comes from the NLOS users. This assumption essentially means that all interfering channels have similar path loss exponent (for example, that of NLOS propagation) and shadowing variance. As such, we express a typical MIMO channel with Tx-Rx distance r in the following form:

$$\mathbf{H} = \sqrt{l(r)}\tilde{\mathbf{H}}, \tag{1}$$

where $\tilde{\mathbf{H}}$ is random and captures the effects of small scale fading, and function $l(r)$ captures the large scale fading which includes both path loss and shadowing. Rician distribution has been used to model the small-scale fading distribution [9]. However, in this paper, we adopt the model in [10] where each element in $\tilde{\mathbf{H}}$ is i.i.d. $\mathcal{CN}(0, 1)$. Consideration of a directional or Rician channel model will be carried out in future work.

The large scale fading function $l(r)$ is modeled as a log-normal random variable multiplied with the pathloss as a function of the distance r as follows:

$$l(r) \;\;=\;\; L_s \beta r^{-\alpha} = L_s l_p(r), \tag{2}$$

where $L_s \sim log(0, \sigma_s)$ is log-normal with standard deviation σ_s dB which captures the shadowing effect; α is the pathloss exponent; and β is the intercept of the pathloss formula. The intercept represents the reference attenuation point that determines the tilt of the path loss model [11].

Based on this channel model, we can express the channels of the direct link between the considered active UE and its BS and the interfering links from other active UEs as follows:

$$\mathbf{H}_0 \;\;=\;\; \sqrt{l\left(\|\mathbf{p}_0\|_2\right)}\tilde{\mathbf{H}}_0, \qquad\qquad \mathbf{H}_k = \sqrt{l\left(\|\mathbf{z}_k\|_2\right)}\tilde{\mathbf{H}}_k, \tag{3}$$

where $\mathbf{p}_0$ and $\mathbf{z}_k$ are vectors representing the 2-D location of the considered active UE and the k^{th} interfering UE in Φ_1 with respect to the origin (i.e. the considered BS). Here $\mathbf{H}_0$ is the direct channel from the considered active UE and $\mathbf{H}_k$ is the channel from the k^{th} interfering UE in Φ_1 to the origin.

III. Interference under MIMO Beamforming

In this section, we present the MIMO beamforming signal model and formulate the out-of-cell interference vector.

A. MIMO Beamforming Signal Model

We now describe the signal model for single-stream MIMO beamforming, i.e., no spatial multiplexing. We model the received signal $\{\mathbf{y}_0 \in \mathbb{C}^{N_{\text{BS}}\times 1}\}$ at the considered BS as

$$\mathbf{y}_0 = \mathbf{H}_0\mathbf{x}_0 + \mathbf{v}_0 + \mathbf{z}_0, \tag{4}$$

where $\mathbf{x}_0 \in \mathbb{C}^{N_{\text{UE}}\times 1}$ is the transmitted signal vector from the considered UE; $\mathbf{z}_0 \in \mathbb{C}^{N_{\text{BS}}\times 1}$ is an i.i.d noise vector distributed as $\mathcal{CN}(\mathbf{0}, \sigma^2 I)$; $\mathbf{H}_0 \in \mathbb{C}^{N_{\text{BS}}\times N_{\text{UE}}}$ is the considered UE-to-BS channel matrix; and $\mathbf{v}_0 \in \mathbb{C}^{N_{\text{BS}}\times 1}$ represents the interference vector received at the considered BS from all interfering UEs.

The transmit signal vector under beamforming from each user can be generally described as

$$\mathbf{x} \;=\; \mathbf{w}\sqrt{P}U. \tag{5}$$

where U is a standard Gaussian signal with zero mean and unit variance; P represents the total power allocated to the active user within a single transmission period; and $\mathbf{w} \in \mathbb{C}^{N_{\text{UE}}\times 1}$ is the unit norm beaforming vector at the active user. We will use this signal model with the appropriate subscript to denote the transmit vector from an active UE, either intended or interfering one.

B. Interference Vector Formulation

For the purpose of modeling network-wide interference, we emphasize that the distribution of interference is independent of the beamforming scheme employed. The only condition is that the transmit and receive beamforming vectors, which are designed for the intended channel, are independent of the interference channels. This condition is realistic in most practical scenarios.

For the interfering active user in the k^{th} cell, denote the interference channel as $\mathbf{H}_k$ and the unit-norm interfering beamforming vectors as $\mathbf{w}_k \in \mathbb{C}^{N_{\text{UE}}\times 1}$. This interfering beamforming vector depends on the direct channel between the k^{th} active UE and its associated BS and hence is independent of the interference channel $\mathbf{H}_k$. Then, we can denote the effective interference vector from the k^{th} interfering active user as

$$\mathbf{h}_k = \mathbf{H}_k\mathbf{w}_k = \sqrt{l\left(\|\mathbf{z}_k\|_2\right)}\tilde{\mathbf{H}}_k\mathbf{w}_k = \sqrt{l\left(\|\mathbf{z}_k\|_2\right)}\tilde{\mathbf{h}}_k, \tag{6}$$

where $\tilde{\mathbf{h}}_k$ is a random vector with i.i.d. elements as $\mathcal{CN}(0, 1)$.

The interference vector can then be expressed as

$$\mathbf{v}_0 = \sum_{k\neq 0}\mathbf{h}_k x_k = \sum_{k\neq 0}\sqrt{P_k l\left(\|\mathbf{z}_k\|_2\right)}\tilde{\mathbf{h}}_k U_k, \tag{7}$$

where the summation is over all interfering users.

For each realization of the UE and BS locations, the out-of-cell interference in Eq. (7) can be modeled as a complex Gaussian random vector with zero mean and covariance matrix $\mathbf{Q}_0$. This matrix is symmetric with diagonal elements representing the interference power at each receiving antenna element, and the off-diagonal elements representing the correlation among interference signals at different antenna elements. In this paper we focus only on modeling the received interference power at each antenna.

IV. Candidate Distributions and Fitness Metric

In this section, we discuss candidate distributions considered for the interference model, the IG and IW distributions. Then, we analytically derive the first moment of the interference power at each antenna element, which is used to match the modeled distribution mean. Finally, we introduce an information-theoretic based metric as a novel way for measuring the model fitness.

A. Candidate Distributions

We consider two candidates for interference modeling: the IG and the IW, both are characterized by two parameters. The IG distribution is used to model nonnegative positively skewed data. It is a two-parameter family of continuous probability distributions which is specified by a shape parameter λ and a scale parameter μ. Given an IG random variable $\gamma_{IG}(\mu,\lambda)$, its probability density function (PDF) is defined as [12]

$$f_{\gamma_{IG}}(t|\mu,\lambda)=\sqrt{\frac{\lambda}{2\pi t^3}}\exp\left\{\frac{-\lambda(t-\mu)^2}{2\mu^2 t}\right\}, t>0, \lambda, \mu>0. \quad (8)$$

The mean and variance of $\gamma_{IG}(\mu,\lambda)$ can be written as [12]

$$\mathbb{E}[\gamma_{IG}]=\mu, \qquad \text{var}[\gamma_{IG}]=\frac{\mu^3}{\lambda}. \quad (9)$$

The IW distribution is used to model data that exhibits a long right tail. It is also specified by two parameters, a shape parameter c and a scale parameter b. Given an IW random variable $\gamma_{IW}(b,c)$, its PDF is defined as [13]

$$f_{\gamma_{IW}}(t|b,c)=\frac{c}{b}\left(\frac{t}{b}\right)^{-c-1}\exp\left\{-\left(\frac{t}{b}\right)^{-c}\right\}, t\geq 0, b, c>0. \quad (10)$$

The mean and variance of $\gamma_{IW}(b,c)$ can be written as [13]

$$\begin{aligned}\mathbb{E}[\gamma_{IW}] &= b\Gamma(1-1/c), \quad \text{if } c>1,\\ \text{var}[\gamma_{IW}] &= b^2\Gamma(1-2/c), \quad \text{if } c>2, \end{aligned} \quad (11)$$

where $\Gamma(t)$ is the Gamma function.

B. Analytical Mean of Interference Power at Each Antenna

The diagonal elements of the interference covariance matrix $\mathbf{Q}_0$ are uncorrelated with the off-diagonal elements. A correlation between any two of these diagonal elements exists and can be determined similarly to the second moment of the off-diagonal elements. However, since this correlation is not strong, we model the diagonal elements of $\mathbf{Q}_0$ as a random vector of independent elements, each element is expressed as

$$q_0 = \sum_{k\neq 0} l\left(\|\mathbf{z}_k\|_2\right) g_k P_k. \quad (12)$$

where g_k are all i.i.d. $\exp(1)$ and represent the power gain of the k^{th} interferer small-scale Rayleigh channel fading.

Lemma 1 below then characterizes the first moment of the out-of-cell interference power q_0 at each receiving antenna.

Lemma 1 (Interference Power Moments)**.** *For network-wide deployment of MIMO transmit beamforming, the out-of-cell interference power at each antenna of the interfered receiver has the following first moment:*

$$\mathbb{E}\left[q_0\right]=\nu P_k R_c^{2-\alpha}, \qquad \nu=\frac{2\pi\lambda_1\beta\mathbb{E}[L_s]}{\alpha-2}. \quad (13)$$

Proof. Omitted due to space, see [14] for details. □

In Sec. V, we use the MLE to estimate the parameters of the mixture distribution. However, to simplify the estimation, we use moment matching first. In which, we match the mean, $\mu_Y = \mathbb{E}\left[q_0\right]$, of the observed data set to that of each of the candidate distributions. This matching results in the scale parameters μ and b for the IG and IW distributions as

$$\mu=\mu_Y, \qquad b=\frac{\mu_Y}{\Gamma(1-1/c)}. \quad (14)$$

Next, we replace these expressions of the scale parameters into the log-likelihood function for each candidate distribution, resulting in a single-parameter log-likelihood functions as

$$\log f_{\gamma_{IG}}(y|\lambda) = \frac{1}{2}\log\lambda-\frac{1}{2}\log 2\pi y^3-\frac{\lambda}{2\mu_Y^2}\frac{(y-\mu_Y)^2}{y}. \quad (15)$$

$$\log f_{\gamma_{IW}}(y|c) = \log\frac{\mu_Y^c c}{\Gamma(1-1/c)^c y^{(c+1)}}-\left(\frac{\mu_Y}{\Gamma(1-1/c)}\right)^c y^{-c}. \quad (16)$$

The derivation and final expressions of the optimal parameters, λ and c are presented in Sec. V-B.

C. KL Divergence Fitness Metric

In order to measure the model goodness or fitness, we use a novel concept from information theory as the relative entropy. The relative entropy of a distribution P with respect to another distribution Q, also called Kullback-Leibler divergence $D_{KL}(P||Q)$, reflects the difference, or distance, between these two probability distributions and is defined as

$$D_{KL}(P||Q)=\int_{-\infty}^{\infty}p(x)\log\frac{p(x)}{q(x)}dx, \quad (17)$$

where $p(x)$ and $q(x)$ denote the PDFs of P and Q, respectively. The distribution P usually represents the true distribution of data or observations.

V. Mixture Interference Model

In this section, we first introduce the PDF of the mixture distribution interference model. Then, we apply MLE to compute the model parameters by developing an EM algorithm.

A. Mixture Interference Model

In our mixture model, the PDF of the interference model is a weighted mixture of both the two-parameter IG and IW distributions. A naive mixture will result in a total of 5 parameters to be estimated, which potentially makes the estimation process complex. We can leverage moment matching, however, to simplify the estimation and represent the mixture model as a 3-parameter distribution as provided in the next theorem.

Theorem 1. *Given a data set Y, the probability density function of the mixture interference model can be written as*

$$f_Y(y|\theta) = w_1 f_{\gamma_{IG}}(y|\lambda)+w_2 f_{\gamma_{IW}}(y|c), \quad (20)$$

1.	**Initialization:**	Initialize $w_j^{(0)}$, $\lambda^{(0)}$, and $c^{(0)}$, $j \in \{1,2\}$, and compute the initial log-likelihood: $$L^{(0)} = \frac{1}{n}\sum_{i=1}^{n} \log\left(w_1^{(0)} f_{\gamma_{IG}}\left(y_i|\lambda^{(0)}\right) + w_2^{(0)} f_{\gamma_{IW}}\left(y_i|c^{(0)}\right)\right). \quad (18)$$	
2.	**E-step:**	Compute $\gamma_{ij}^{(m)}$ as given in Eq. (23) and $n_j^{(m)} = \sum_{i=1}^{n}\gamma_{ij}^{(m)}$.	
3.	**M-step:**	Compute the new estimate for $w_j^{(m+1)}$, $\lambda^{(m+1)}$, and $c^{(m+1)}$, $j \in \{1,2\}$ as given in Eqs. (26), and (27).	
4.	**Convergence check:**	Compute the new log-likelihood function $$L^{(m+1)} = \frac{1}{n}\sum_{i=1}^{n}\log\left(\sum_{j=1}^{2} w_1^{(m+1)} f_{\gamma_{IG}}\left(y_i|\lambda^{(m+1)}\right) + w_2^{(m+1)} f_{\gamma_{IW}}\left(y_i|c^{(m+1)}\right)\right). \quad (19)$$	
	Return to 2 if $|L^{(m+1)} - L^{(m)}| \geq \delta$ for a preset threshold δ; **Otherwise** end the algorithm.		

Algorithm 1. Mixture MLE Model EM Algorithm

where $\{w_1, w_2 : w_1 + w_2 = 1\}$ are the weight parameters for mixing the IG and IW distributions and $f_{\gamma_{IG}}(y|\lambda)$ and $f_{\gamma_{IW}}(y|c)$ are as given in Eqs. (15) *and* (16) *with the scale parameters obtained through moment matching first as in* (14)*. This mixture model is a 3-parameter distribution with the parameter set as $\theta = \{w_1, \lambda, c\}$.*

B. MLE Model Parameters Estimation

We have three parameters to be estimated in the mixture model which makes the EM algorithm [15] an appealing approach. Here, we identify the complete data X as the observed data Y plus some hidden data Z, i.e. $X = (Y, Z)$, where Y is the observed set of points that we model by a weighted IG and IW distribution and $Z \in \{1, 2\}$ is a discrete random variable representing the assignment of each data point to the two candidate distributions. Then, we can define the PDF of the complete data $X \in \mathbb{R}^+$ as

$$f_X(x|\theta) = f_X(Y = y, Z = i|\theta) = w_i \phi_i(y|\theta_i), \quad (21)$$

where $\phi_1(y|\theta_1) = f_{\gamma_{IG}}(y|\lambda)$ and $\phi_2(y|\theta_2) = f_{\gamma_{IW}}(y|c)$.

Next, we present the algorithm for estimating the parameters of the MLE interference models. Given n observations, in order to develop the EM algorithm, we use Proposition 1 below [16] which can be easily verified using the independence assumption, the fact that the observed data $Y = f(X)$ is a deterministic function of the complete data X for some function f, and Bayes' rule.

Proposition 1. [16] *Let the complete data X consist of n i.i.d. samples: $X_1, X_2, ..., X_n$, which satisfies $f(X|\theta) = \prod_{i=1}^{n} f(X_i|\theta)$ for all $\theta \in \Theta$, and let $y_i = f(x_i)$, $i = 1, 2, ..., n$, then*

$$Q(\theta|\theta^{(m)}) = \sum_{i=1}^{n} Q_i(\theta|\theta^{(m)}), \quad (22)$$

We also define a responsibility function $\gamma_{ij}^{(m)}$, which represents our guess at the m^{th} iteration of the probability that the i^{th} sample belongs to the j^{th} distribution component, as

$$\gamma_{ij}^{(m)} = p\left(Z_i = j|Y_i = y_i, \theta^{(m)}\right) = \frac{w_j^{(m)}\phi_j(y_i|\theta_j^{(m)})}{\sum_{l=1}^{2} w_l^{(m)}\phi_l(y_i|\theta_l^{(m)})}, i \in \{1, ..., n\}, j \in \{1, 2\} \quad (23)$$

where θ_1 and θ_2 represent the shape parameters λ and c.

An EM algorithm includes two steps: the E-step to calculate the conditional expectation of the log likelihood of the complete data, and the M-step to maximize this conditional expectation function. Lemma 2 provides the final expression for the EM algorithm Q-function in the E-step.

Lemma 2. *Let the complete data X consist of n i.i.d. samples: $X_1, X_2, ..., X_n$, which satisfies $f(X|\theta) = \prod_{i=1}^{n} f(X_i|\theta)$ for all $\theta \in \Theta$, and let $y_i = T(x_i)$, $i = 1, 2, ..., n$, then we have*

$$\begin{aligned} Q(\theta|\theta^{(m)}) &= \sum_{j=1}^{2} n_j^{(m)} \log w_j + \frac{1}{2} n_1^{(m)} \log\lambda - \frac{\lambda}{2\mu_Y^2}\sum_{i=1}^{n}\gamma_{i1}^{(m)}\frac{(y_i - \mu_Y)^2}{y_i} \\ &+ c n_2^{(m)} \log\mu_Y - c n_2^{(m)} \log\Gamma(1 - 1/c) + n_2^{(m)} \log c \\ &- c\sum_{i=1}^{n}\gamma_{i2}^{(m)} \log y_i - \left[\frac{\mu_Y}{\Gamma(1-1/c)}\right]^c \sum_{i=1}^{n}\gamma_{i2}^{(m)} y_i^{-c}, \end{aligned} \quad (24)$$

where $n_j^{(m)} = \sum_{i=1}^{n}\gamma_{ij}^{(m)}, j \in \{1, 2\}$; and $\theta = \{w_1, \lambda, c\}$.

Proof. See Appendix A for details. □

In the M-step, to update our estimate of the parameter θ, we solve the following problem for the optimal θ^*:

$$\arg\max_{\theta}\ Q(\theta|\theta^{(m)}) \ \text{ s.t. } \sum_{j=1}^{2} w_j = 1, w_j \geq 0, j \in \{1, 2\}. \quad (25)$$

The optimal θ^* is then found as in Theorem 2.

Theorem 2. *The optimal new estimate of $\theta^{(m+1)} = \left\{w_1^{(m+1)}, \lambda^{(m+1)}, c^{(m+1)}\right\}$ that maximizes the Q-function in* (24) *at the m^{th} iteration are determined as follows:*

$$w_j^{(m+1)} = \frac{n_j^{(m)}}{n}, j \in \{1, 2\}, \lambda^{(m+1)} = \frac{n_1^{(m)}\mu_Y^2}{\sum_{i=1}^{n}\gamma_{i1}^{(m)}\frac{(y_i - \mu_Y)^2}{y_i}}, \quad (26)$$

and $c^{(m+1)}$ is obtained by numerically solving the equation:

$$\begin{aligned} 0 &= n_2^{(m)}\left[\log\frac{\mu_Y}{\Gamma(1-\frac{1}{c})} + \frac{1}{c}\left[1 - \psi\left(1 - \frac{1}{c}\right)\right]\right] - \sum_{i=1}^{n}\gamma_{i2}^{(m)}\log y_i \\ &+ \left[\frac{\mu_Y}{\Gamma\left(1-\frac{1}{c}\right)}\right]^c \sum_{i=1}^{n}\gamma_{i2}^{(m)} y_i^{-c}\left[\log\frac{y_i\Gamma(1-\frac{1}{c})}{\mu_Y} + \frac{\psi(1-\frac{1}{c})}{c}\right] \end{aligned} \quad (27)$$

Proof. See Appendix B for details. □

Algorithm 1 summarizes the EM algorithm for the mixture distribution interference model.

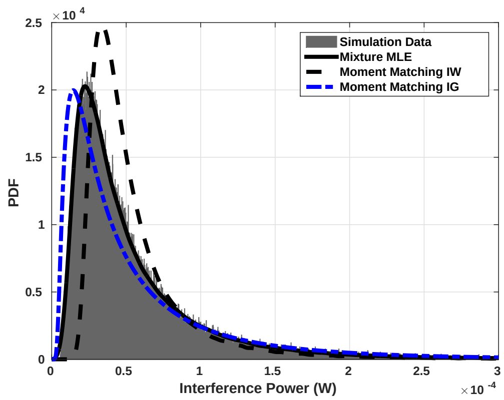

Fig. 1: Sample PDF of interference models at $\alpha = 3.5$, $\sigma_{SF} = 4$ dB, ($P_{max} = 30$ dBm, $\delta = 10^{-6}$).

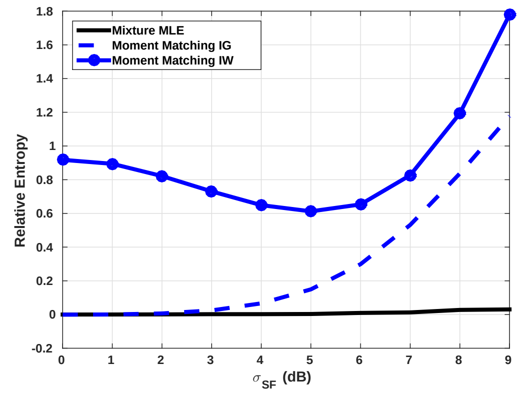

Fig. 2: Relative entropy of different interference power models at $\alpha = 3.5$, ($P_{max} = 30$ dBm, $\delta = 10^{-6}$).

VI. Numerical Performance Results

In this section, we use simulation data obtained from the stochastic geometry based network setting in Sec. II-A to numerically verify the validity of the proposed mixture interference model. For the simulations setting, we consider a typical cell of radius $R_c = 150$ m, which is related to the active UEs density λ_1 as $\lambda_1 = 0.25\eta R_c^{-2}$, where $\eta = 1$ in the typical case and can be varied to examine the effect of user density on performance.

For reference, we compare our mixture model with the moment matching interference models in which a simple matching between the first two interference power moments and those of a candidate distribution with a known two-parameter PDF is used [17]. We then apply these interference models to evaluate the outage performance of a system employing MIMO joint transmit-receive dominant mode beamforming. In this case, the transmit and receive beamforming vectors are the two left and right singular vectors corresponding to the dominant mode of the channel matrix between the respective transmitter and receiver.

A. Mixture MLE Interference Model Evaluation

In Fig. 1, we show a sample interference power distribution at one antenna element for $\alpha = 3.5$ and $\sigma_{SF} = 4$ dB to visually confirm the match, or rather the mismatch, of the moment-matched IW interference distribution. In Fig. 2, we plot the relative entropy versus shadowing standard deviation σ_{SF} at a sample path loss exponent value of $\alpha = 3.5$. The result shows that IG approximates the simulated interference distribution well at low values of shadowing variance and diverges at higher values. The IW distribution, on the other hand, does not show a good fit at the simulated path loss value. We further see that our proposed mixture model outperforms both the individual IG and IW models. This result aligns with our expectation since the PDF of the mixture model is a combination of both the IG and IW distributions. Further, the objective of MLE is to minimize the log difference between the entire set of experimental data and the modeled distribution while the MM techniques only matches the first two moments.

Consistently in all our extensive simulations and testing, the mixture MLE interference model provides the best fit. This model, however, requires an iterative algorithm to fit the parameters. The MM models are attractive from the simplicity point of view since the model parameters can be determined analytically without optimization [14], [17]. For this reason, we identify in Fig. 3 the regions of α and σ_{SF} parameters in which the IG MM model has a reasonably good fit. In this range of channel parameters, the relative entropy of the IG MM model is within 1% of the Mixture model.

B. Rate CDF and Outage Performance

We now apply the proposed interference models to evaluate the cellular system performance. In particular, we evaluate the transmission rate cumulative distribution function (CDF) of a cell edge user at a distance of 145 m from the BS. All user equipments and BSs are equipped with 2 antennas, and the uplink maximum transmission power is 30 dBm. We use the dominant mode beamforming and assume a noise variance of $\sigma^2 = -124$ dBm/Hz and a path loss intercept $\beta = -72.3$ dB.

In Fig. 4, we compare the user transmission rate CDFs based on IG MM and mixture interference models to the simulated data at $\alpha = 3$ and $\sigma_{SF} = 9$ dB. At such large shadowing, only the mixture model remains accurate while the IG MM diverges significantly, in agreement with the results in Fig. 3.

VII. Conclusion

We have introduced a new interference model as a mixture of the Inverse Gaussian and Inverse Weibull distributions. We design an EM algorithm to estimate the parameters of this mixture model and compare it to moment matching models of the individual distributions. An information theoretic based metric is exploited to measure the relative distance between the developed interference models and the simulated data. The mixture model, despite being computationally more complex in parameter estimation than moment matching, offers the most accurate fit. As a future work, we plan to directly fit the optimal mixture model parameters to simple functions of the mmWave propagation characteristics, including path loss and shadowing parameters.

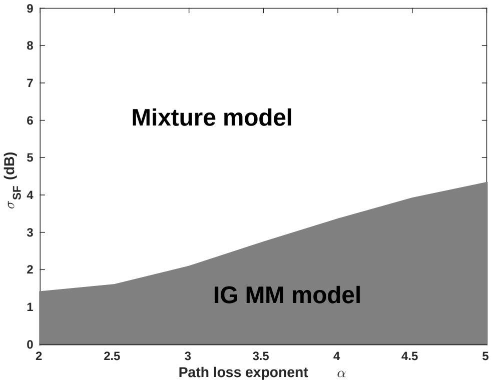

Fig. 3: Regions of shadowing standard deviation and path loss exponent for the IG MM model to fit within 1% of the mixture MLE model ($P_{max} = 30$ dBm, $\eta = 1$).

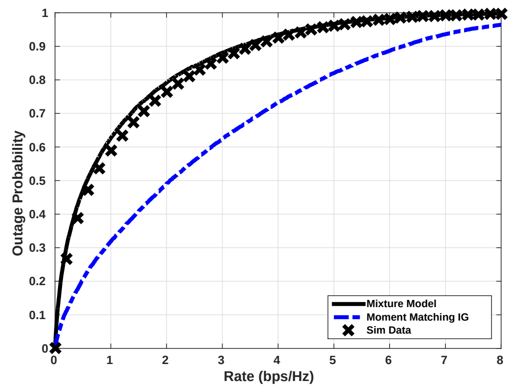

Fig. 4: User transmission rate CDF comparison among simulation, IG MM model, and mixture MLE model. System settings: $P_{max} = 30$ dBm, $\eta = 1$, $\alpha = 3$, $\sigma_{SF} = 9$ dB.

Appendix A: Proof of Lemma 2

We start by writing the expression for the EM algorithm Q-function using Proposition 1 as

$$\begin{aligned} Q(\theta|\theta^{(m)}) &= \sum_{i=1}^{n} \mathbb{E}_{X_i|Y_i,\theta^{(m)}} \left[\log f_{X_i}(x_i|\theta)\right] \\ &= \sum_{i=1}^{n}\sum_{j=1}^{2} \gamma_{ij}^{(m)} \log w_j + \sum_{i=1}^{n} \gamma_{i1}^{(m)} \log \phi_1(y_i|\theta_1) \\ &+ \sum_{i=1}^{n} \gamma_{i2}^{(m)} \log \phi_2(y_i|\theta_2). \end{aligned} \quad (28)$$

Using the probability density functions of IG and IW distributions in Eqs. (8) and (10) along with moment matching first, we can write $\log\phi_1(y_i|\lambda)$ and $\log\phi_2(y_i|c)$ as in Eqs. (15) and (16), respectively. Then, substituting into Eq. (28), replacing $\sum_{i=1}^{n} \gamma_{ij}^{(m)}$ by $n_j^{(m)}$, and dropping constant terms, we get the final expression in Eq. (24).

Appendix B: Proof of Theorem 2

We first solve for w_j using the method of Lagrange multipliers as follows.

$$J(w,\nu) = \sum_{j=1}^{2} n_j^{(m)} \log w_j + \nu \left(1 - \sum_{j=1}^{2} w_j\right). \quad (29)$$

Equating to zero the partial derivatives of $J(w,\nu)$ w.r.t. w_j and substituting into the constraint $\sum_{j=1}^{2} w_j = 1$, we get the weight w_j update as

$$w_j^{(m+1)} = \frac{n_j^{(m)}}{\sum_{j=1}^{k} n_j^{(m)}} = \frac{n_j^{(m)}}{\sum_{j=1}^{k}\sum_{i=1}^{n} \gamma_{ij}^{(m)}} = \frac{n_j^{(m)}}{n}, j \in \{1,2\}. \quad (30)$$

Then, we can find the update of λ_{IW} by simply taking the partial derivative of the Q-function w.r.t. λ_{IW} as

$$\frac{\partial Q\left(\theta|\theta^{(m)}\right)}{\partial \lambda} = \frac{1}{2\lambda} n_1^{(m)} - \frac{1}{2\mu_Y^2} \sum_{i=1}^{n} \gamma_{i1}^{(m)} \frac{(y_i - \mu_Y)^2}{y_i} = 0. \quad (31)$$

Solving for $\lambda^{(m+1)}$ results in Eq. (26).

Similarly, we can find the update of c by taking the partial derivative of the Q-function w.r.t. c and equating to zero, we obtain the equation in (27).

References

[1] S. Rajagopal, S. Abu-Surra, and M. Malmirchegini, "Channel feasibility for outdoor non-line-of-sight mmwave mobile communication," in *IEEE Vehicular Tech. Conf. (VTC Fall)*, Sept 2012, pp. 1–6.

[2] W. Roh, J.-Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: theoretical feasibility and prototype results," *IEEE Comm. Magazine*, vol. 52, no. 2, pp. 106–113, Feb. 2014.

[3] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Comm. Magazine*, vol. 52, no. 2, pp. 74–80, Feb. 2014.

[4] T. Bai, V. Desai, and R. W. Heath, "Millimeter wave cellular channel models for system evaluation," in *Proc. IEEE ICNC*, 2014, pp. 178–182.

[5] T. Bai, A. Alkhateeb, and R. Heath, "Coverage and capacity of millimeter-wave cellular networks," *IEEE Comm. Magazine*, vol. 52, no. 9, pp. 70–77, Sept. 2014.

[6] M. Kountouris and N. Pappas, "Approximating the interference distribution in large wireless networks," in *2014 11th Proc. IEEE Int. Symp. on Wireless Comm. Systems (ISWCS)*, pp. 80–84.

[7] M. R. Akdeniz, Y. Liu, M. K. Samimi, S. Sun, S. Rangan, T. S. Rappaport, and E. Erkip, "Millimeter wave channel modeling and cellular capacity evaluation," *IEEE Journal on Sel. Areas in Comm.*, vol. 32, no. 6, pp. 1164–1179, 2014.

[8] T. Bai and R. Heath, "Coverage and rate analysis for millimeter-wave cellular networks," *IEEE Trans. on Wireless Comm.*, vol. 14, no. 2, pp. 1100–1114, Feb 2015.

[9] M. K. Samimi, S. Sun, and T. S. Rappaport, "Mimo channel modeling and capacity analysis for 5g millimeter-wave wireless systems," in *2016 10th European Conference on Antennas and Propagation (EuCAP)*, April 2016, pp. 1–5.

[10] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[11] G. R. MacCartney, J. Zhang, S. Nie, and T. S. Rappaport, "Path loss models for 5G millimeter wave propagation channels in urban microcells," in *2013 Proc. IEEE Global Comm. Conf. (GLOBECOM)*.

[12] M. C. Tweedie, "Statistical properties of inverse gaussian distributions. I." *The Annals of Math. Stat.*, pp. 362–377, 1957.

[13] H. Rinne, *The Weibull distribution: a handbook*. CRC Press, 2008.

[14] H. Elkotby and M. Vu, "Interference modeling for 5G millimeter wave beamforming," *Submitted to IEEE Trans. on Wireless Comm.*

[15] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38, 1977.

[16] Y. Chen and M. R. Gupta, "Em demystified: An expectation-maximization tutorial," *University of Washington*, 2010.

[17] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Comm. Surveys Tutorials*, vol. 15, no. 3, pp. 996–1019, 2013.